基于两种相似度矩阵的专利引文耦合方法识别研究前沿*

——以脑机接口为例

高 楠 傅俊英 赵蕴华
(中国科学技术信息研究所 北京 100038)

摘要: 【目的】探索基于两种相似度矩阵的专利引文耦合分析方法以识别研究前沿。【方法】基于原始观测值和余弦距离两种相似度算法, 建立专利相似度矩阵, 利用社会网络分析得到研究前沿簇, 并进行簇类命名, 从而得到研究前沿。并利用 Innography 数据库的脑机接口领域专利对以上方法进行案例研究。【结果】发现两种相似度矩阵中, 基于原始观测值算法得到 6 个研究前沿簇, 涉及 6 类 BCI 研究内容; 基于余弦距离算法得到 9 个研究前沿簇, 涉及 8 类 BCI 研究内容, 两者的 FID 重合率均为 43%。【局限】本文侧重于两种算法的结果, 即研究前沿数量、重合度和内容进行比较, 缺少对于算法本身特性的比较。【结论】基于这两种相似度算法的引文耦合法均可识别出领域的研究前沿, 余弦距离相似度算法能识别出更多数量的研究前沿, 且比基于原始观测值相似度算法的识别结果更全面。

关键词: 研究前沿 专利 耦合分析 脑机接口 相似度算法
分类号: G255.53 R318.6

1 引 言

研究前沿(Research Fronts, RF)最早由 De Solla Price[1]于 1965 年提出, 即一组高被引文献, 用于描述领域内的瞬态特征和动态变化, 是对领域已有主题进行的评价和分析。1974 年, Small 等[2]首先提出使用共被引分析来识别活跃领域的研究前沿。而基于耦合分析识别研究前沿的定义也于 1994 年由 Persson[3]提出, 即对一簇引用共同参考文献形成了耦合关系的相似文献, 其施引文献构成了研究前沿, 而被引的参考文献称为研究前沿的知识基础。Persson 还结合文献耦合分析、作者共被引分析和文献共被引分析等方法研究了 1986 年–1990 年 JASIS 中的文章, 分析了情报学领域的研究前沿及其与知识基础的关系。Glanzel 等[4]也提出耦合分析是识别研究前沿的新的有效方法。之后, Morris 等[5]将研究前沿进一步描述为一簇趋于引用固定的、不随时间变化的基础文献, 即文献耦合集, 以获得一个领域的整体发展趋势、可视化结构和动态变化。同时, 采用一种创新的研究前沿分析和展现的时间线方法, 即基于文献耦合聚类, 并引入连续的时间轴, 以展示各个研究前沿的出现增长和减弱消失的时间点, 进而得到研究前沿的演进变化, 以及潜在新研究前沿等信息。Yang 等[6]利用耦合分析对 1978 年–2006 年囊状纤维化症领域进行分析, 得到 20 项研究前沿, 特别是与骨质密度相关的研究前沿。张婷[7]利用耦合分析识别了科学传播领域的研究前沿。另外, 有学者结合耦合分析和其他分析方法对研究前沿进行识别。如 Jarneving[8]结合耦合分析和完全连接聚类分析识别了严重呼吸道症候群领域的研究前沿。Chen 等[9]利用专利耦合分析结合自然语言处理技术对美国和全球在

通讯作者: 高楠, ORCID: 0000-0002-9152-9545, E-mail: 1183621517@qq.com。
*本文系中央级公益性科研院所基本科研业务费专项资金项目“未来学”(项目编号:XK2015-2)的研究成果之一。

智能电网行业的研究前沿进行比较。因此，在文献计量学领域，研究前沿既可以通过论文数据的识别表征科学基础前沿，也可以通过专利数据的识别表征创新技术前沿。

研究表明，文献耦合方法具有较好的时效性，因为论文发表后即可获得参考文献数据，而不需要其被引用数据，故最近发表的文献也能被聚成一类[10]，能够比较快速地反映领域研究前沿的进展变化。但其也不可避免地存在所有基于引用关系的文献计量方法的限制：如引用行为的偏好，引用内容不同侧重点不同，不同领域的文献被人为聚在一起，分析样本的代表性问题，忽视部分重要的文献，以及数据库的问题等[11]。

目前文献计量学领域对于研究前沿的识别分析，多以论文作为数据源，较少涉及专利数据的分析。世界知识产权组织公布的一组数据显示[12]，专利说明书中含有 90%-95%的研发成果，其中 85%以上的技术将不再出现于其他技术文献中，且要比其他载体早公开一到两年，因此专利数据涵盖更为丰富的技术信息，适于领域的研究前沿分析。但专利数据具有其特殊性，在实际分析中要针对实际情况做灵活的处理。本文将在已有基于论文耦合方法识别研究前沿的基础上，以专利为数据源，对专利型数据做相应处理后，进行研究前沿的识别。

2 研究方法

2.1 数据源的选择

专利数据存在专利家族的特殊性，一个专利家族的同族专利指拥有相同优先权专利，在不同国家或组织多次申请、多次公布或审核批准的一系列内容相同或相似的专利文献[13]，表征相同或相似的技术。故本文对专利家族，而非单个专利进行研究。ProQuest Dialog 公司的 Innography 专利信息检索和分析平台有专利家族号(Family ID, FID)字段，同族专利的家族成员拥有相同的 FID，代表一个技术点，因而采用该数据库作为数据源进行研究前沿分析。数据库中的专利后向引用指标(Backward Citations, BC)即该专利对其他专利的引用情况，可用于评估发明的新颖性及其技术基础，在本研究中用于耦合分析[14]。

2.2 专利相似度矩阵的建立

将同一 FID 所有家族成员的后引数据合并，使得每个 FID 仅出现一次。但是，在合并后引数据时，对于一个专利家族多个成员都引用的相同专利，存在两种计数方式，产生出两种相似度计算方法，进而衍生出两种耦合相似度矩阵生成方法。

(1) 基于原始观测值(Observed Value, OV)的相似度计算方法，即在合并后引数据时，对于相同的后引专利全部归一，无论后引专利被专利家族中几个专利成员引用，都只计数 1 次，专利对间的相似度即为两个专利家族相同后引专利的数量。

(2) 基于余弦距离(Cosine Distance, CD)的相似度计算方法，是将专利家族合并后的后引数据转化成空间向量，然后计算两两向量间夹角的余弦值，得到余弦距离(也称为余弦相似度[15])。余弦距离用向量间的夹角来衡量相似度，重视向量在方向上的差异。两个 n 维向量 A [A_1, A_2, …, A_n]和 B [B_1, B_2, …, B_n]间的余弦距离为：

$$\text{Distance} = \cos\theta = \frac{\sum_{i=1}^{n}(A_i \times B_i)}{\sqrt{\sum_{i=1}^{n}(A_i)^2} \times \sqrt{\sum_{i=1}^{n}(B_i)^2}} = \frac{A \cdot B}{\|A\|\|B\|} \quad (1)$$

第一种方法反映的是专利对间绝对的共现次数，可以真实地反映共现次数，但该计算方法的缺点是在家族所有的相同后引专利归一化后，不能反映它们在家族中所占的比重。比如，某一项后引专利在多个家族成员的施引数据中出现，则反映了该后引专利对这一专利家族存在引用偏好，但归一化处理后，这种引用偏好无法体现。第二种方法反映的是专利对间相对的共现次数，这种将专利的引用情况向量化，以向量的夹角大小反映专利对间的相似度，可以保留专利家族的引用偏好，但缺点是无法还原现实的共现次数。两种方法各有利弊，互为补充，因此本文采用两种相似度计算方法建立耦合相似度矩阵。

相似度矩阵的行和列均为 FID，根据专利后向引用指标和两种相似度算法分别计算专利家族对间的相似度，遍历所有专利家族对后，即可分别建立基于原始观测值的耦合相似度矩阵(OV-BCA)和基于余弦距离的耦合相似度矩阵(CD-BCA)。

2.3 研究前沿识别

根据张嘉彬[11]的研究，利用社会网络分析法对两种耦合相似度矩阵进行聚类，连续调节阈值直到能获

得比较清晰的聚类结果，并且聚类数量控制在 15 个以内，所得簇类即研究前沿。同时，将包含三个及以上的专利家族聚类算作一个研究前沿。

最后，对识别出的研究前沿进行命名，依据其共同的合作专利分类号(Cooperative Patent Classification, CPC)以及摘要和权利要求项的内容综合运用人工解读、切词技术和专家智慧对研究前沿簇进行命名。

3　脑机接口领域实证分析

脑科学领域的脑机接口技术(Brain Computer Interface, BCI)指在无外周神经系统和肌肉组织参与的条件下，通过计算机等电子设备采集转化大脑活动信号为输出控制信号，进而与外界环境进行交流。BCI 开辟了一种全新的模式，给人类提供了一种可根据不同情境的大脑活动来操控电脑或者通信设备进行活动的可能[16]，为用意念或思维控制外部设备提供了可行手段[17]。其涉及计算机科学、神经生理学和康复医学等领域[18]，被国际多个权威机构选为21世纪的研究前沿和热点之一[19]，因此识别脑机接口领域的研究前沿意义重大。

3.1　数据源

本研究在 Innography 专利信息检索和分析平台检索到脑机接口技术领域相关专利共 3 546 件，检索时间范围为 2006 年 1 月 1 日–2015 年 5 月 30 日，检索日期为: 2015 年 5 月 30 日。进行数据清洗剔除无关专利，最终得到相关专利 3 543 件，涉及专利家族 1 848 个。

检索策略由 4 部分组成:

①BCI 的各种变形表达: @(abstract, pclaims, title) ("Brain machine interface*") OR ("brain response interface*") OR ("Brain and Computer Interface*") OR ("direct neural interface") OR ("Brain Computer Interface*")@*

②BCI 技术分类的表达: (@(abstract, pclaims, title) ("brain activity") OR ("neurophysiological signal*") OR ("cortical neuronalactivity") OR ("cortical neuronal action potential*")) AND (@(abstract, pclaims, title) ("Steady-State Visual Evoked Potential*") OR ("Event Related Potential or Evoked Potential*") OR (" Slow Cortical Potential ") OR ("Visual Evoked Potential ") OR ("P300 evoked potential*") OR ("P300") OR ("event related desynchronization") OR ("Event Related synchronization") OR ("mental task") OR ("spontaneous rhythm*"))@*

③BCI 技术的国际专利分类号: @meta CPC_G06F0003015000@*

④BCI 技术相关的 CPC 再利用 BCI 技术分类进行限定: ((@meta CPC_A61B0005044800) OR (@meta CPC_A61B0005047600) OR (@meta CPC_A61B0005040010) OR (@meta CPC_A61B0005004200)) AND (@(abstract, pclaims, title) (Signal Acquisition) OR (Feature Extraction) OR (Translation Algorithm) OR (Device Command*) OR ("Brain Computer Interface*") OR ("Brain machine interface*") OR ("brain response interface*") OR ("Brain and Computer Interface*") OR ("direct neural interface"))@*

3.2　基于原始观测值相似度算法

根据第 2 节所述方法，利用 VBA 编程计算专利对间相似度，得到 1848×1848 的脑机接口领域的耦合相似度矩阵。识别聚类簇得到图 1 所示结果，图中每个圆圈代表一个技术点，以专利家族号表示，线的粗细表示 FID 间联系的紧密程度。基于原始观测值的耦合分析法(OV-BCA)，调节观测值的阈值为 10，聚类出的研究前沿簇达到 6 个，涵盖 35 个专利家族，共 116 件专利。最后，对研究前沿簇进行命名，得到表 1 中的命名结果。图 1 中右侧的聚类簇所含专利家族号最多，达到 21 个，进一步细分可得到两个研究前沿: B5 神经响应测量的市场化应用与用户评价; B6 利用植入式神经刺激器进行神经系统调制，治疗神经系统疾病。

表 1 揭示的是零散的前沿技术点，为便于宏观分析，对脑机接口研究前沿按照 BCI 的分类、组成和应用三方面共 13 个类别对零散的前沿技术点进行归类，具体包括: 植入式 BCI 和非植入式 BCI。非植入 BCI 又可细分为基于诱发响应的BCI、基于自发响应的BCI 和 fNIRS-BCI(Functional Near-Infrared Spectroscopy-BCI); 信号采集、信号处理和控制器; 辅助神经系统类疾病的诊断或治疗、脑功能研究、交流和通信、运动与转移、游戏娱乐和康复训练。按上述 13 个类别对 B1 到 B6 涉及的内容进行归类: 非植入式 BCI(B1)，植入式 BCI(B6)，信号采集(B4、B5)，信号处理(B2)，辅助诊断和治疗(B3)，脑功能研究(B6)，共 6 类。

反映了 BCI 在近 10 年的两个前沿重点: 一个是专注于 BCI 的市场化应用，且引入了用户的喜好评价，使得 BCI 的设计更加人性化; 另一个是从之前的以无创性非植入式的 EEG 作为 BCI 的主要信号源，发展到微创性的植入式 BCI 研究，这也是 BCI 发展的必然趋势。2013 年举行的第 5 届国际脑机接口会议上对于 BCI 未来发展方向中也提到相同内容[20]: 相比于非植入式 BCI，植入式 BCI 展现出其对大脑活动信号控制

更好的质量和速度，而成为新的研究聚焦点，及未来对于 BCI 设计用户体验友好性、易用性的不断追求。国际脑机接口会议每 3 年举行一次，于 2013 年举行了第 5 届脑机接口国际会议，其所展示的研究前沿内容具备权威性，且最能反映近 10 年，乃至未来 1-3 年的重要发展方向。

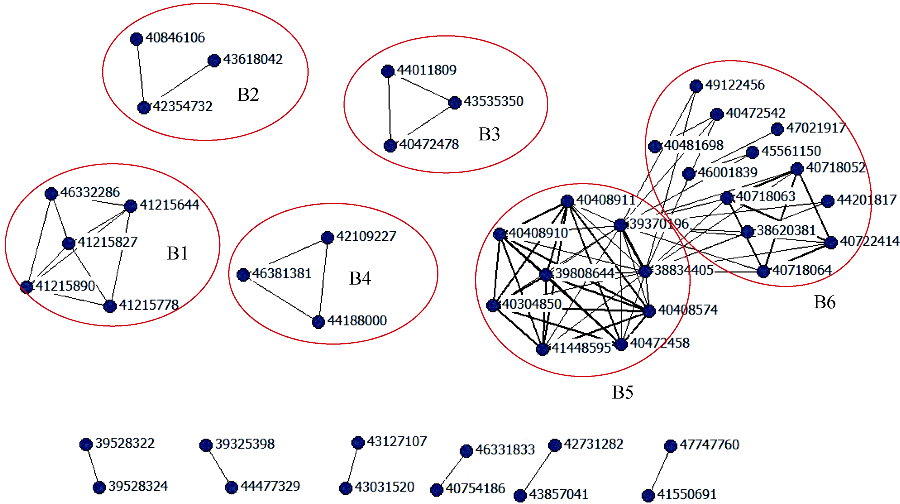


图 1　基于 OV-BCA 识别 BCI 研究前沿

表 1　基于 OV-BCA 的研究前沿识别结果

时间	阈值	序号	研究前沿	聚类中专利		数量		所占比例(%)	
				平均优先权年	平均公布年	专利家族	专利	专利家族	专利
2006-2015	10	B1	用于视觉刺激电路的生物活性剂	2008	2011	5	9	14.3	7.8
		B2	使用小波变换对脑电图的波形信号进行处理	2009	2011	3	13	8.6	11.2
		B3	用于调节神经系统活动或治疗的磁疗法，如脑电同步化磁震荡治疗	2008	2011	3	16	8.6	13.8
		B4	基于任务相关刺激产生的自发脑电的检测、测量或记录	2008	2012	3	6	8.6	5.2
		B5	神经响应测量(如脑磁图(MEG)、脑电图(EEG)、皮电反应(GSR)、心电图(EKG)、眼动电图(EOG)，眼动跟踪和面部情绪编码测量等)的市场化应用与用户评价	2007	2011	9	32	25.7	27.5
		B6	利用植入式神经刺激器进行神经系统调制(如电、热、磁调制)，治疗神经系统疾病(如神经系统紊乱)	2008	2011	12	40	34.2	34.5

3.3　基于余弦距离相似度算法

根据第 2 节所述方法，利用 TDA(Thomson Data Analyzer)工具计算专利对间余弦相似度，得到 1848×1848 的脑机接口领域的耦合相似度矩阵。调节余弦距离的阈值为 0.53，识别聚类簇得到图 2 所示结果，基于余弦距离的耦合分析法(CD-BCA)，得到 9 个研究前沿簇，涵盖 35 个专利家族，共 59 件专利。对研究前沿簇进行命名，得到表 2 中的命名结果。图 2 中所占比例最大的是 b3 神经响应测量的市场化应用与用户评价，达到 20%。b1-b9 涉及的内容包括：非植入式 BCI(b5)，植入式 BCI(b6)，信号采集(b2、b3、b7、b9)，信号处理(b8)，辅助诊断和治疗(b4)，脑功能研究(b6)，运动与转移(b1)，康复训练(b7)，共 8 类。

识别出的 B5 和 b3 所包含的研究前沿，二者 FID 高度一致，且均在近 10 年内占据很大比例，一方面表明这就是该时间窗内的重要研究前沿，另一方面说明对于相关度较强且引用次数较高的多个专利家族，用不同的相似度计算方法都能将其识别出来。此外，B1 和 b7 二者所含 FID 也高度一致。

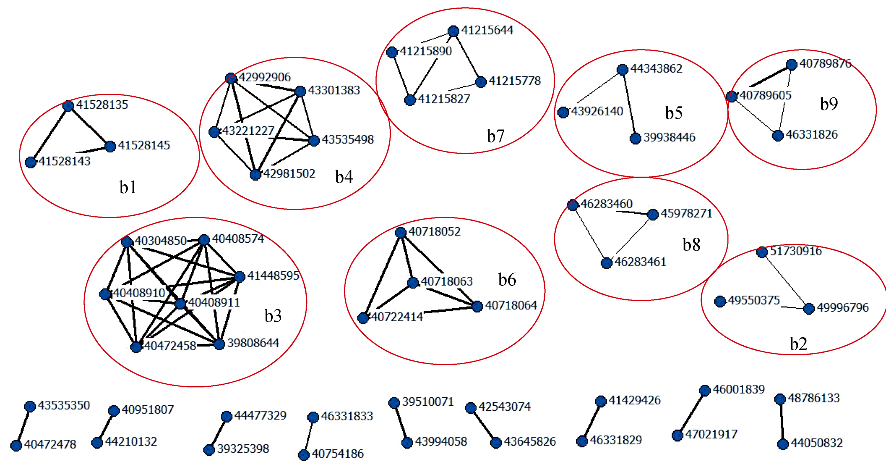


图 2　基于 CD-BCA 识别 BCI 研究前沿

表 2　基于 CD-BCA 的研究前沿识别结果

时间	阈值	序号	研究前沿	聚类中专利		数量		所占比例(%)	
				平均优先权年	平均公布年	专利家族	专利	专利家族	专利
		b1	以神经系统活动为输入的控制系统，如控制车辆	2008	2010	3	3	8.6	5.1
		b2	用于脑的成像设备，如核磁共振，光学断层扫描或阻抗断层扫描装置	2012	2014	3	3	8.6	5.1
		b3	神经响应测量(如脑磁图(MEG)、脑电图(EEG)、皮电反应(GSR)、心电图(EKG)、眼动电图(EOG)，眼动跟踪和面部情绪编码测量等)的市场化应用与用户评价	2007	2012	7	18	20	30.5
		b4	计算机辅助医疗诊断或治疗的医疗保健及康复的系统和方法	2009	2010	5	5	14.2	8.5
2006-2015	0.53	b5	基于视觉诱发电位的控制系统，例如视觉刺激生成器，基于虚拟现实的脑机接口，动眼追踪等	2009	2012	3	6	8.6	10.2
		b6	利用植入式神经刺激器进行神经活动调制，如热调制、电调制、磁调制等	2008	2010	4	10	11.4	16.9
		b7	用于视觉刺激电路的生物活性剂	2008	2011	4	7	11.4	11.8
		b8	基于小波分析技术的脑电图运动区功能定位系统	2012	2012	3	3	8.6	5.1
		b9	利用生理活性数据(如脑电数据、脑血氧水平等)反映人的精神、心理状态，如脑电图，正电子发射断层扫描，磁共振成像，功能磁共振成像，功能性近红外成像，或脑磁图	2008	2009	3	4	8.6	6.8

仔细比较表 1 和表 2 内容可以发现，CD-BCA 的识别内容几乎涵盖了 OV-BCA 识别出的所有内容，包括：非植入式 BCI、植入式 BCI、信号采集、信号处理、辅助诊断和治疗、脑功能研究，此外还识别出 OV-BCA 未识别出的 BCI 涉类，如运动与转移及康复训练。因而，CD-BCA 能识别出 OV-BCA 识别不出的新的前沿簇，其前沿识别能力更强。

3.4　结果对比

汇总基于原始观测值的耦合分析法(OV-BCA)和基于余弦距离的耦合分析法(CD-BCA) 的研究前沿识

别结果，如表 3 所示。可以看出，基于余弦距离的相似度算法能识别出的更多数量的研究前沿簇。

计算 OV-BCA 和 CD-BCA 的研究前沿重合率，发现两者的 FID 重合率都为 43%，其中，重合的 FID 尤其能够代表已引发学科共同体广泛关注的高质量研究，如表 4 所示。同时，由 3.4 节已知 CD-BCA 能识别出 OV-BCA 识别不出的前沿簇，其前沿识别能力更强。

表 3　基于 OV-BCA 和 CD-BCA 所得研究前沿情况

比较项	相似度计算方法	数值
阈值	OV-BCA	10
	CD-BCA	0.53
前沿个数(个)	OV-BCA	6
	CD-BCA	9
前沿所包含的FID 数量(个)	OV-BCA	35
	CD-BCA	35

表 4　OV-BCA 和 CD-BCA 的研究前沿重合率

重合 FID 数量	前沿数量(个)		前沿所包含的 FID 总数(个)		FID 重合率	
	OV-BCA	CD-BCA	OV-BCA	CD-BCA	OV-BCA	CD-BCA
15	6	9	35	35	0.43	0.43

4　结　语

本文采用引文耦合分析法，利用专利数据源，基于两种相似度矩阵，分析了脑机接口领域近 10 年的研究前沿情况。同时对基于原始观测值和基于余弦距离的两种相似度算法进行了对比研究。

引文耦合分析识别研究前沿，最关键的一步是耦合相似度矩阵的建立，之后可通过聚类识别出所探求的前沿结果。因为专利数据具有专利家族的特殊性，而一个专利家族对应一个技术点，因而相比于单篇专利更适合于研究前沿中技术点的识别。然而，两个专利家族间相似度的计算，是两类专利文献集间相似度的计算，在合并同一 FID 所有家族成员的后引数据后，由于对相同后引数据的计数有两种方式，因此，必然导致两种相似度计算方式的使用，并分别获得两种耦合相似度矩阵及不同的研究前沿。结果表明，这两种方法都可以识别出脑机接口领域的研究前沿情况；同时，结果表明基于余弦距离相似度算法的耦合方法的前沿识别能力更强，所得前沿数量更多，且识别结果更细致。

但是，本文仅侧重于对两种相似度算法结果的比较，即对研究前沿数量、重合度和内容进行比较，对于算法本身特性的比较并没有展开，存在一定的欠缺。另外，仅选取一个领域进行研究，并且是按照专利家族建立的矩阵，数据量不够丰富，后续将选取多个领域进行案例研究，以进一步验证两种算法的适应性。同时，因为专利的重要引用关系除了耦合关系，还有共被引关系。未来将在本耦合方法研究的基础上，探索基于共被引方法进行研究前沿识别的方法，并比较两者结果的异同，以期对研究前沿识别这一关键问题获得较为全面科学的探索。

参考文献：

[1] De Solla Price D J. Networks of Scientific Papers[J]. Science, 1965, 149(3683): 510-515.

[2] Small H G, Griffith B C. The Structure of Scientific Literatures I: Identifying and Graphing Specialties [J]. Science Studies, 1974, 4(1): 17-40.

[3] Persson O. The Intellectual Base and Research Fronts of JASIS 1986-1990 [J]. Journal of the American Society for Information Science, 1994, 45(1): 31-38.

[4] Glanzel W, Czerwon H J. A New Methodological Approach to Bibliographic Coupling and Its Application to the National, Regional and Institutional Level [J]. Scientometrics, 1996, 37(2): 195-221.

[5] Morris S A, Yen G, Wu Z, et al. Time Line Visualization of Research Fronts [J]. Journal of the American Society for Information Science and Technology, 2003, 54(5): 413-422.

[6] Yang L, Morris S A, Barden E M. Mapping Institutions and Their Weak Ties in a Research Specialty: A Case Study of Cystic Fibrosis Body Composition Research [J]. Scientometrics, 2009, 79(2): 421-434.

[7] 张婷. 时间线和地形式可视化图谱: 科学传播研究前沿演进趋势分析[J]. 情报学报, 2009, 28(6): 923-928. (Zhang Ting. Timeline and Landscape: A Case Study of Visualizing the Evolution of Science Communication Research Front[J]. Journal of the China Society for Scientific & Technical Information, 2009, 28(6): 923-928.)

[8] Jarneving B. The Outbreak of SARS Mirrored by

Bibliometric Mapping: Combining Bibliographic Coupling with the Complete Link Cluster Method[J]. Library and Information Science Research, 2007, 11(11): 11-36.

[9] Chen H, Huang M H, Chen D Z, et al. Detecting the Temporal Gaps of Technology Fronts: A Case Study of Smart Grid Field[J]. Technological Forecasting & Social Change, 2012, 79(9): 1705-1719.

[10] Hicks D. Limitations of Co-Citation Analysis as a Tool for Science Policy[J]. Social Studies of Science, 1987, 17(2): 295-316.

[11] 张嘉彬. 以书目耦合及共被引探讨不同引用区间之研究前沿: 以 OLED 领域为例[D]. 台北: 台湾大学, 2011. (Zhang Jiabin. The Study of Different Citation Windows of Research Fronts in OLED—Using Bibliographic Coupling and Co-Citation Methods [D]. Taipei: Taiwan University, 2011.)

[12] 张强. 基于专利计量的专利实施许可实证研究[D]. 重庆: 西南政法大学, 2012. (Zhang Qiang. Empirical Research of Patent Licensing Based on Patentometrics [D]. Chongqing: Southwest University of Political Science and Law, 2012.)

[13] 吴琳, 魏星, 霍翠婷. 基于 Web 的专利双语语料自动获取研究及实现——以 esp@cenet 数据库为例[J]. 现代图书情报技术, 2009(9): 57-63. (Wu Lin, Wei Xing, Huo Cuiting. Research and Implement of Automatic Patent Bilingual Corpus Extraction from Web——Taking esp@cenet as an Example [J]. New Technology of Library and Information Service, 2009(9): 57-63.)

[14] 经济合作与发展组织编著. 专利统计手册[M]. 高昌林等译. 北京: 科学技术文献出版社, 2013: 3. (Organization for Economic Co-operation and Development. Patent Statistics Manua [M]. Translated by Gao Changlin, et al. Beijing: Scientific and Technical Documention Press, 2013: 3.)

[15] 刘妍. 基于 Lucene 的余弦距离检测文档相似度方法的研究[J]. 信息系统工程, 2014(4): 129-130, 142. (Liu Yan. The Research of Document Similarity Detection Based on Lucene Cosine Distance Method [J]. China CIO News, 2014(4): 129-130, 142.)

[16] Wang B, Wong C M, Wan F, et al. Comparison of Different Classification Methods for EEG-based Brain Computer Interfaces: A Case Study [C]. In: Proceedings of the 2009 IEEE International Conference on Information and Automation. 2009.

[17] 杨建, 刘进, 吴明曦, 等. 脑机技术发展及其对军事领域的影响[J]. 国防科技, 2013, 34(6): 19-23. (Yang Jian, Liu Ming, Wu Mingxi, et al. Brain-Computer Technology Development and Its Impact on the Military [J]. National Defense Science & Technology, 2013, 34(6): 19-23.)

[18] 尧德中, 刘铁军, 雷旭, 等. 基于脑电的脑–机接口: 关键技术和应用前景[J]. 电子科技大学学报, 2009, 38(5): 550-554. (Yao Dezhong, Liu Tiejun, Lei Xu, et al. Electroencephalogram Based Brain-Computer Interface: Key Techniques and Application Prospect [J]. Journal of University of Electronic Science and Technology of China, 2009, 38(5): 550-554.)

[19] Mason S G, Bashshati A, Fatourechi M, et a1. A Comprehensive Survey of Brain Interface Technology Designs [J]. Annals of Biomedical Engineering, 2007, 35(2): 137-169.

[20] Huggins J E, Guger C, Allison B, et al. Workshops of the Fifth International Crain-Computer Interface Meeting: Defining the Future [J]. Brain-Computer Interfaces, Taylor & Francis, 2014, 1(1): 27-49.

作者贡献声明:

高楠: 共同设计研究方案, 针对专利数据的专利家族特点提出两种相似度计算方法, 进行研究前沿耦合识别方法和脑机接口领域研究进展的文献调研, 制定专利检索策略, 获取、处理与分析数据, 撰写及修订论文;

傅俊英: 提出用专利引文耦合方法识别研究前沿的研究思路, 共同设计研究方案, 进行研究前沿耦合识别方法的文献调研, 并对论文最终版本提出修改意见;

赵蕴华: 共同设计研究方案, 并对论文最终版本提出修改意见。

利益冲突声明:

所有作者声明不存在利益冲突关系。

支撑数据:

支撑数据见期刊网络版 http://www.infotech.ac.cn。

[1] 高楠, 傅俊英, 赵蕴华. 06-15CD-BCA.xlsx. 基于余弦距离算法的专利相似度矩阵.

[2] 高楠, 傅俊英, 赵蕴华. 06-15OV-BCA.xlsx. 基于原始观测值的专利相似度矩阵.

收稿日期: 2015-09-29
收修改稿日期: 2015-11-08

Identifying Research Trends Based on Patent Bibliographic Coupling: Case Study of Brain-Computer Interface

Gao Nan　Fu Junying　Zhao Yunhua
(Institute of Scientific & Technical Information of China, Beijing 100038, China)

Abstract: [**Objective**] This study aims to identify the research trends (RT) based on patent bibliographic coupling method with the help of similarity algorithms. [**Methods**] We first established two types of patent similarity matrixes with two similarity algorithms - observed value (OV-BCA) and cosine distance (CD-BCA). We then used social network analysis to get the RT of Brain-Computer Interface (BCI) Patents. [**Results**] Six BCI research trend clusters were retrieved by OV-BCA algorithm, while CD-BCA algorithm got nine RT clusters. The two algorithms' family ID coincidence rates were 43%. [**Limitations**] We focused on the comparison of results, including number, content and coincidence degree. More research is needed to study the characteristics of these algorithms. [**Conclusions**] RT can be retrieved by bibliographic coupling method with the help of the proposed algorithms. Specifically, the cosine distance algorithm can find more detailed research trends than the observed value algorithm.

Keywords: Research trends　Patent　Bibliographic coupling analysis　Brian computer interface　Similarity algorithm